1a5s
Evolutionary trace report by report maker
April 18, 2010

CONTENTS

1 Introduction

2 Chain 1a5sA
  2.1 P00929 overview
  2.2 Multiple sequence alignment for 1a5sA
  2.3 Residue ranking in 1a5sA
  2.4 Top ranking residues in 1a5sA and their position on the structure
     2.4.1 Clustering of residues at 25% coverage.
     2.4.2 Overlap with known functional surfaces at 25% coverage.

3 Chain 1a5sB
  3.1 Q5PD17 overview
  3.2 Multiple sequence alignment for 1a5sB
  3.3 Residue ranking in 1a5sB
  3.4 Top ranking residues in 1a5sB and their position on the structure
     3.4.1 Clustering of residues at 25% coverage.
     3.4.2 Overlap with known functional surfaces at 25% coverage.
     3.4.3 Possible novel functional surfaces at 25% coverage.

4 Notes on using trace results
  4.1 Coverage

5 Appendix
  5.1 File formats
  5.2 Color schemes used
  5.3 Credits
     5.3.1 Alistat
     5.3.2 CE
     5.3.3 DSSP
     5.3.4 HSSP
     5.3.5 LaTex
     5.3.6 Muscle
     5.3.7 Pymol
  5.4 Note about ET Viewer
  5.5 Citing this work
  5.6 About report maker
  5.7 Attachments

1 INTRODUCTION

1 From the original Protein Data Bank entry (PDB id 1a5s):
Title: Crystal structure of wild-type tryptophan synthase complexed with 5-fluoroindole propanol phosphate and L-ser bound as amino acrylate to the beta site
Compound: Mol id: 1; molecule: tryptophan synthase (alpha chain);
chain: a; ec: 4.2.1.20; engineered: yes; mol id: 2; molecule: tryptophan synthase (beta chain); chain: b; ec: 4.2.1.20; engineered: yes
Organism, scientific name: Salmonella Typhimurium;
1a5s contains unique chains 1a5sA (258 residues) and 1a5sB (387 residues)

2 CHAIN 1A5SA

2.1 P00929 overview
From SwissProt, id P00929, 89% identical to 1a5sA:
Description: Tryptophan synthase alpha chain (EC 4.2.1.20).
Organism, scientific name: Salmonella typhimurium.
Taxonomy: Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.
Function: The alpha subunit is responsible for the aldol cleavage of indoleglycerol phosphate to indole and glyceraldehyde 3- phosphate.
Catalytic activity: L-serine + 1-(indol-3-yl)glycerol 3-phosphate = L-tryptophan + glyceraldehyde 3-phosphate.
Pathway: Amino-acid biosynthesis; L-tryptophan biosynthesis; L-tryptophan from chorismate: step 5 [final step].
Fig. 1. Residues 2-130 in 1a5sA colored by their relative importance. (See Appendix, Fig.16, for the coloring scheme.)

Subunit: Tetramer of two alpha and two beta chains.
Similarity: Belongs to the trpA family.
About: This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

2.2 Multiple sequence alignment for 1a5sA
For the chain 1a5sA, the alignment 1a5sA.msf (attached) with 342 sequences was used. The alignment was downloaded from the HSSP database, and fragments shorter than 75% of the query as well as duplicate sequences were removed. It can be found in the attachment to this report, under the name of 1a5sA.msf. Its statistics, from the aIstat program are the following:

<table>
<thead>
<tr>
<th>Format</th>
<th>MSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sequences</td>
<td>342</td>
</tr>
<tr>
<td>Total number of residues</td>
<td>84,283</td>
</tr>
<tr>
<td>Smallest</td>
<td>181</td>
</tr>
<tr>
<td>Largest</td>
<td>258</td>
</tr>
<tr>
<td>Average length</td>
<td>246.4</td>
</tr>
<tr>
<td>Alignment length</td>
<td>258</td>
</tr>
<tr>
<td>Average identity</td>
<td>42%</td>
</tr>
<tr>
<td>Most related pair</td>
<td>99%</td>
</tr>
<tr>
<td>Most unrelated pair</td>
<td>24%</td>
</tr>
<tr>
<td>Most distant seq</td>
<td>44%</td>
</tr>
</tbody>
</table>

Furthermore, 1% of residues show as conserved in this alignment. The alignment consists of 4% eukaryotic (2% fungi, 2% plantae), 23% prokaryotic, and 1% archaean sequences. (Descriptions of some sequences were not readily available.) The file containing the sequence descriptions can be found in the attachment, under the name of 1a5sA.descr.

2.3 Residue ranking in 1a5sA
The 1a5sA sequence is shown in Figs. 1–2, with each residue colored according to its estimated importance. The full listing of residues in 1a5sA can be found in the file called 1a5sA.ranks_sorted in the attachment.

2.4 Top ranking residues in 1a5sA and their position on the structure
In the following we consider residues ranking among top 25% of residues in the protein. Figure 3 shows residues in 1a5sA colored by their importance: bright red and yellow indicate more conserved/important residues (see Appendix for the coloring scheme). A Pymol script for producing this figure can be found in the attachment.

Fig. 2. Residues 131-267 in 1a5sA colored by their relative importance. (See Appendix, Fig.16, for the coloring scheme.)

2.4.1 Clustering of residues at 25% coverage. Fig. 4 shows the top 25% of all residues, this time colored according to clusters they belong to. The clusters in Fig.4 are composed of the residues listed in Table 1.

<table>
<thead>
<tr>
<th>cluster</th>
<th>color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
</table>
|         | red   | 57   | 22,26,27,28,49,51,53,54,55
|         |       |      | 56,57,59,60,61,64,65,67,70
|         |       |      | 71,72,100,102,104,105,125
|         |       |      | 127,129,130,132,134,135,151
|         |       |      | 153,155,156,157,158,162,172
|         |       |      | 173,175,177,178,181,182,183
|         |       |      | 184,207,211,212,213,214,232

continued in next column
Fig. 4. Residues in 1a5sA, colored according to the cluster they belong to: red, followed by blue and yellow are the largest clusters (see Appendix for the coloring scheme). Clockwise: front, back, top and bottom views. The corresponding Pymol script is attached.

Table 1. Clusters of top ranking residues in 1a5sA.

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue</td>
<td>4</td>
<td>234,235,236,238</td>
</tr>
<tr>
<td>yellow</td>
<td>2</td>
<td>110,114</td>
</tr>
</tbody>
</table>

2.4.2 Overlap with known functional surfaces at 25% coverage.
The name of the ligand is composed of the source PDB identifier and the heteroatom name used in that file.

**FIP binding site.** Table 2 lists the top 25% of residues at the interface with 1a5sAFIP270 (fip). The following table (Table 3) suggests possible disruptive replacements for these residues (see Section 4.6).

### Table 2.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>G</td>
<td>G(92).</td>
<td>0.08</td>
<td>1/1</td>
<td>4.78</td>
</tr>
<tr>
<td>213</td>
<td>G</td>
<td>G(92).</td>
<td>0.08</td>
<td>9/9</td>
<td>2.74</td>
</tr>
<tr>
<td>64</td>
<td>I</td>
<td>I(96)N V(2).</td>
<td>0.09</td>
<td>7/0</td>
<td>4.50</td>
</tr>
<tr>
<td>212</td>
<td>F</td>
<td>F(92).</td>
<td>0.09</td>
<td>46/16</td>
<td>2.99</td>
</tr>
<tr>
<td>59</td>
<td>A</td>
<td>A(93). L(4) S(1) T(1).</td>
<td>0.11</td>
<td>10/7</td>
<td>3.78</td>
</tr>
<tr>
<td>234</td>
<td>G</td>
<td>G(90).</td>
<td>0.11</td>
<td>16/16</td>
<td>2.94</td>
</tr>
<tr>
<td>235</td>
<td>S</td>
<td>S(90).</td>
<td>0.12</td>
<td>19/11</td>
<td>2.57</td>
</tr>
<tr>
<td>177</td>
<td>L</td>
<td>L(29). V(66) A I(2) Q</td>
<td>0.14</td>
<td>1/0</td>
<td>4.88</td>
</tr>
<tr>
<td>129</td>
<td>A</td>
<td>A(37) C V(25) P(35) TL</td>
<td>0.15</td>
<td>4/1</td>
<td>3.93</td>
</tr>
<tr>
<td>100</td>
<td>L</td>
<td>L(41). M(35) F(21) YX</td>
<td>0.16</td>
<td>29/0</td>
<td>3.67</td>
</tr>
<tr>
<td>127</td>
<td>L</td>
<td>L(50). I(33) V(15) T</td>
<td>0.17</td>
<td>3/0</td>
<td>4.38</td>
</tr>
<tr>
<td>153</td>
<td>I</td>
<td>I(37). L(57) M(2) F(1) V</td>
<td>0.17</td>
<td>4/0</td>
<td>2.92</td>
</tr>
<tr>
<td>214</td>
<td>I</td>
<td>I(86).</td>
<td>0.19</td>
<td>3/3</td>
<td>4.28</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>F(74).</td>
<td>0.20</td>
<td>22/0</td>
<td>3.56</td>
</tr>
<tr>
<td>232</td>
<td>I</td>
<td>I(71).</td>
<td>0.20</td>
<td>6/2</td>
<td>3.90</td>
</tr>
<tr>
<td>236</td>
<td>A</td>
<td>A(84) R(2) H V(1) ECQ TI</td>
<td>0.21</td>
<td>1/1</td>
<td>4.52</td>
</tr>
</tbody>
</table>

Table 2. continued

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>Y</td>
<td>Y(95) C(4) S</td>
<td>0.06</td>
<td>3/0</td>
<td>4.04</td>
</tr>
<tr>
<td>184</td>
<td>G</td>
<td>G(95). S(4)</td>
<td>0.05</td>
<td>11/11</td>
<td>3.27</td>
</tr>
<tr>
<td>60</td>
<td>D</td>
<td>D(94) E(5)</td>
<td>0.08</td>
<td>14/0</td>
<td>2.67</td>
</tr>
</tbody>
</table>

Table 2. The top 25% of residues in 1a5sA at the interface with FIP (Field names: res: residue number in the PDB entry; type: amino acid type; substs: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

continued in next column
Table 3. List of disruptive mutations for the top 25% of residues in 1a5sA, that are at the interface with FIP.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>E</td>
<td>(FW)(YH)(VCAG)(T)</td>
</tr>
<tr>
<td>175</td>
<td>Y</td>
<td>(K)(QM)(NEVLAPIR)(D)</td>
</tr>
<tr>
<td>183</td>
<td>T</td>
<td>(KR)(FQMWH)(NELPI)(D)</td>
</tr>
<tr>
<td>184</td>
<td>G</td>
<td>(KR)(E)(FQMWH)(D)</td>
</tr>
<tr>
<td>102</td>
<td>Y</td>
<td>(K)(M)(Q)(R)</td>
</tr>
<tr>
<td>60</td>
<td>D</td>
<td>(R)(FWH)(YVCAG)(K)</td>
</tr>
<tr>
<td>211</td>
<td>G</td>
<td>(KER)(FQMWH)(NLPI)(Y)</td>
</tr>
<tr>
<td>213</td>
<td>G</td>
<td>(KER)(FQMWH)(NLPI)(Y)</td>
</tr>
<tr>
<td>64</td>
<td>I</td>
<td>(Y)(R)(H)(T)</td>
</tr>
<tr>
<td>212</td>
<td>F</td>
<td>(K)(E)(Q)(D)</td>
</tr>
<tr>
<td>59</td>
<td>A</td>
<td>(R)(K)(Y)(E)</td>
</tr>
<tr>
<td>234</td>
<td>G</td>
<td>(KER)(Q)(HD)(M)</td>
</tr>
<tr>
<td>235</td>
<td>S</td>
<td>(R)(K)(H)(FQW)</td>
</tr>
<tr>
<td>177</td>
<td>L</td>
<td>(Y)(R)(H)(T)</td>
</tr>
<tr>
<td>129</td>
<td>A</td>
<td>(R)(K)(Y)(E)</td>
</tr>
<tr>
<td>100</td>
<td>L</td>
<td>(R)(Y)(T)(K)</td>
</tr>
<tr>
<td>127</td>
<td>L</td>
<td>(R)(Y)(H)(K)</td>
</tr>
<tr>
<td>153</td>
<td>I</td>
<td>(R)(Y)(T)(H)</td>
</tr>
<tr>
<td>214</td>
<td>I</td>
<td>(Y)(R)(H)(T)</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>(K)(E)(Q)(D)</td>
</tr>
<tr>
<td>232</td>
<td>I</td>
<td>(Y)(R)(H)(T)</td>
</tr>
<tr>
<td>236</td>
<td>A</td>
<td>(Y)(E)(R)(K)</td>
</tr>
</tbody>
</table>

Figure 5 shows residues in 1a5sA colored by their importance, at the interface with 1a5sAFIP270.

**Interface with 1a5sB.** Table 4 lists the top 25% of residues at the interface with 1a5sB. The following table (Table 5) suggests possible disruptive replacements for these residues (see Section 4.6).

Table 4.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's</th>
<th>cvg</th>
<th>noc/ dist (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>D</td>
<td>D(100)</td>
<td>0.02</td>
<td>47/35 2.98</td>
</tr>
<tr>
<td>181</td>
<td>G</td>
<td>G(99)A</td>
<td>0.02</td>
<td>27/27 3.55</td>
</tr>
<tr>
<td>57</td>
<td>P</td>
<td>P(99)H</td>
<td>0.03</td>
<td>15/12 3.05</td>
</tr>
<tr>
<td>56</td>
<td>D</td>
<td>D(99)E</td>
<td>0.04</td>
<td>77/14 2.75</td>
</tr>
<tr>
<td>132</td>
<td>P</td>
<td>P(98)ST</td>
<td>0.05</td>
<td>48/15 3.14</td>
</tr>
<tr>
<td>104</td>
<td>N</td>
<td>N(98)GS</td>
<td>0.06</td>
<td>64/14 2.57</td>
</tr>
<tr>
<td>53</td>
<td>P</td>
<td>P(97)</td>
<td>0.07</td>
<td>6/6 3.03</td>
</tr>
<tr>
<td>60</td>
<td>D</td>
<td>D(94)</td>
<td>0.08</td>
<td>10/10 3.09</td>
</tr>
<tr>
<td>54</td>
<td>F</td>
<td>F(78)</td>
<td>0.09</td>
<td>31/9 3.48</td>
</tr>
<tr>
<td>65</td>
<td>Q</td>
<td>Q(97)</td>
<td>0.10</td>
<td>27/0 3.28</td>
</tr>
<tr>
<td>59</td>
<td>A</td>
<td>A(93)</td>
<td>0.11</td>
<td>4/3 4.35</td>
</tr>
<tr>
<td>156</td>
<td>P</td>
<td>P(96)QT</td>
<td>0.12</td>
<td>2/0 4.05</td>
</tr>
<tr>
<td>162</td>
<td>L</td>
<td>L(12)</td>
<td>0.12</td>
<td>10/0 3.77</td>
</tr>
<tr>
<td>72</td>
<td>F</td>
<td>F(85)</td>
<td>0.14</td>
<td>12/0 3.59</td>
</tr>
<tr>
<td>177</td>
<td>L</td>
<td>L(29)</td>
<td>0.14</td>
<td>4/0 4.36</td>
</tr>
<tr>
<td>129</td>
<td>A</td>
<td>A(37)C</td>
<td>0.15</td>
<td>13/13 3.47</td>
</tr>
<tr>
<td>55</td>
<td>S</td>
<td>S(85)</td>
<td>0.16</td>
<td>37/9 2.76</td>
</tr>
<tr>
<td>153</td>
<td>I</td>
<td>I(37)</td>
<td>0.17</td>
<td>8/6 3.92</td>
</tr>
<tr>
<td>135</td>
<td>E</td>
<td>E(86)</td>
<td>0.19</td>
<td>75/2 2.54</td>
</tr>
<tr>
<td>182</td>
<td>V</td>
<td>V(87)</td>
<td>0.21</td>
<td>11/1 3.86</td>
</tr>
</tbody>
</table>

continued in next column

Fig. 5. Residues in 1a5sA, at the interface with FIP, colored by their relative importance. The ligand (FIP) is colored green. Atoms further than 30 Å away from the geometric center of the ligand, as well as on the line of sight to the ligand were removed. (See Appendix for the coloring scheme for the protein chain 1a5sA.)
Table 4. The top 25% of residues in 1a5sA at the interface with 1a5sB. 
(Field names: res: residue number in the PDB entry; type: amino acid type; 
subs: substitutions seen in the alignment; with the percentage of each type 
in the bracket; noc/bb: number of contacts with the ligand, with the number of 
contacts realized through backbone atoms given in the bracket; dist: distance 
of closest approach to the ligand.)

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subs's ( % )</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>155</td>
<td>P</td>
<td>T(4) I(8)</td>
<td>0.22</td>
<td>13/0</td>
<td>4.00</td>
</tr>
<tr>
<td>134</td>
<td>E</td>
<td>K(1) L(4)</td>
<td>0.23</td>
<td>22/0</td>
<td>2.72</td>
</tr>
<tr>
<td>105</td>
<td>L</td>
<td>L(39) P(52)T</td>
<td>0.24</td>
<td>14/1</td>
<td>3.67</td>
</tr>
<tr>
<td>157</td>
<td>N</td>
<td>N(28) S(4)</td>
<td>0.25</td>
<td>29/1</td>
<td>3.33</td>
</tr>
</tbody>
</table>

Table 5. List of disruptive mutations for the top 25% of residues in 
1a5sA, that are at the interface with 1a5sB.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>D</td>
<td>(R) (FWH) (KYVCAG) (TQM)</td>
</tr>
<tr>
<td>181</td>
<td>G</td>
<td>(KER) (QHD) (FYMW) (N)</td>
</tr>
<tr>
<td>57</td>
<td>P</td>
<td>(TYR) (E) (SKCG) (QHD)</td>
</tr>
<tr>
<td>56</td>
<td>D</td>
<td>(R) (FWH) (YVCAG) (K)</td>
</tr>
<tr>
<td>132</td>
<td>P</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>104</td>
<td>N</td>
<td>(Y) (FWH) (R) (E)</td>
</tr>
<tr>
<td>53</td>
<td>P</td>
<td>(YR) (H) (TKE) (SQCDG)</td>
</tr>
<tr>
<td>60</td>
<td>D</td>
<td>(R) (FWH) (YVCAG) (K)</td>
</tr>
<tr>
<td>54</td>
<td>F</td>
<td>(KE) (Q) (D) (T)</td>
</tr>
<tr>
<td>65</td>
<td>Q</td>
<td>(Y) (H) (FW) (T)</td>
</tr>
<tr>
<td>59</td>
<td>A</td>
<td>(R) (K) (Y) (E)</td>
</tr>
<tr>
<td>156</td>
<td>P</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>162</td>
<td>L</td>
<td>(R) (Y) (H) (T)</td>
</tr>
<tr>
<td>72</td>
<td>F</td>
<td>(KE) (T) (R) (QD)</td>
</tr>
<tr>
<td>177</td>
<td>L</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>129</td>
<td>A</td>
<td>(R) (K) (Y) (E)</td>
</tr>
</tbody>
</table>

Table 5. continued

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>S</td>
<td>(KR) (FQMWH) (NELPI) (Y)</td>
</tr>
<tr>
<td>153</td>
<td>I</td>
<td>(R) (Y) (T) (H)</td>
</tr>
<tr>
<td>135</td>
<td>E</td>
<td>(H) (FW) (R) (Y)</td>
</tr>
<tr>
<td>182</td>
<td>V</td>
<td>(R) (K) (Y) (E)</td>
</tr>
<tr>
<td>155</td>
<td>P</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>134</td>
<td>E</td>
<td>(H) (Y) (FW) (R)</td>
</tr>
<tr>
<td>105</td>
<td>L</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>157</td>
<td>N</td>
<td>(Y) (H) (FW) (R)</td>
</tr>
</tbody>
</table>

Table 4. continued

Fig. 6. Residues in 1a5sA, at the interface with 1a5sB, colored by their relative importance. 1a5sB is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5sA.)

Figure 6 shows residues in 1a5sA colored by their importance, at the interface with 1a5sB.

3 CHAIN 1A5SB

3.1 Q5PD17 overview

From SwissProt, id Q5PD17, 93% identical to 1a5sB: 
Description: Tryptophan synthase beta chain. 
Organism, scientific name: Salmonella paratyphi-a. 
Taxonomy: Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

3.2 Multiple sequence alignment for 1a5sB

For the chain 1a5sB, the alignment 1a5sB.msf (attached) with 511 sequences was used. The alignment was downloaded from the HSSP database, and fragments shorter than 75% of the query as well as
3.3 Residue ranking in 1a5sB

The 1a5sB sequence is shown in Figs. 7–8, with each residue colored according to its estimated importance. The full listing of residues in 1a5sB can be found in the file called 1a5sB.ranks sorted in the attachment.

3.4 Top ranking residues in 1a5sB and their position on the structure

In the following we consider residues ranking among top 25% of residues in the protein. Figure 9 shows residues in 1a5sB colored by their importance: bright red and yellow indicate more conserved/important residues (see Appendix for the coloring scheme). A Pymol script for producing this figure can be found in the attachment.

3.4.1 Clustering of residues at 25% coverage. Fig. 10 shows the top 25% of all residues, this time colored according to clusters they belong to. The clusters in Fig. 10 are composed of the residues listed in Table 6.

### Table 6.

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue</td>
<td>3</td>
<td>193,196,198</td>
</tr>
</tbody>
</table>

**continued in next column**
Fig. 10. Residues in 1a5sB, colored according to the cluster they belong to: red, followed by blue and yellow are the largest clusters (see Appendix for the coloring scheme). Clockwise: front, back, top and bottom views. The corresponding Pymol script is attached.

Table 6. Clusters of top ranking residues in 1a5sB.

<table>
<thead>
<tr>
<th>cluster</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>yellow</td>
<td>2</td>
<td>256,269</td>
</tr>
</tbody>
</table>

3.4.2 Overlap with known functional surfaces at 25% coverage.

The name of the ligand is composed of the source PDB identifier and the heteroatom name used in that file.

**PLP binding site.** Table 7 lists the top 25% of residues at the interface with 1a5sCPLP901 (plp). The following table (Table 8) suggests possible disruptive replacements for these residues (see Section 4.6).

Table 7. The top 25% of residues in 1a5sB at the interface with PLP.(Field names: res: residue number in the PDB entry; type: amino acid type; subst’s: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>304</td>
<td>L</td>
<td>(99)M.</td>
<td>0.16</td>
<td>4/1</td>
<td>3.88</td>
</tr>
<tr>
<td>114</td>
<td>Q</td>
<td>(95)</td>
<td>0.17</td>
<td>8/1</td>
<td>3.29</td>
</tr>
<tr>
<td>382</td>
<td>K</td>
<td>(97)</td>
<td>0.21</td>
<td>1/0</td>
<td>4.59</td>
</tr>
<tr>
<td>378</td>
<td>G</td>
<td>(97)</td>
<td>0.23</td>
<td>3/3</td>
<td>4.02</td>
</tr>
</tbody>
</table>

Table 8. List of disruptive mutations for the top 25% of residues in 1a5sB, that are at the interface with PLP.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>86</td>
<td>H</td>
<td>(E) (QMD) (SNKVCNALPG) (YR)</td>
</tr>
<tr>
<td>234</td>
<td>G</td>
<td>(KER) (FMWHD) (NYLPI) (NVA)</td>
</tr>
<tr>
<td>169</td>
<td>G</td>
<td>(KER) (FMWHD) (Y) (NLPI)</td>
</tr>
<tr>
<td>232</td>
<td>G</td>
<td>(R) (K) (FWH) (EQM)</td>
</tr>
<tr>
<td>236</td>
<td>N</td>
<td>(Y) (FWH) (TR) (VCAG)</td>
</tr>
<tr>
<td>87</td>
<td>K</td>
<td>(Y) (T) (FW) (SCG)</td>
</tr>
<tr>
<td>233</td>
<td>G</td>
<td>(E) (D) (FMW) (YQLPHIR)</td>
</tr>
<tr>
<td>348</td>
<td>A</td>
<td>(KYER) (QHD) (N) (FTMW)</td>
</tr>
<tr>
<td>235</td>
<td>S</td>
<td>(KR) (FMWHD) (NELPI)</td>
</tr>
<tr>
<td>303</td>
<td>G</td>
<td>(KR) (E) (FMWHD) (Q)</td>
</tr>
<tr>
<td>350</td>
<td>E</td>
<td>(FWH) (Y) (VCARG) (T)</td>
</tr>
<tr>
<td>304</td>
<td>L</td>
<td>(Y) (R) (T) (H)</td>
</tr>
<tr>
<td>114</td>
<td>Q</td>
<td>(Y) (H) (FW) (T)</td>
</tr>
<tr>
<td>382</td>
<td>K</td>
<td>(Y) (FTW) (SVCAAG) (HD)</td>
</tr>
<tr>
<td>378</td>
<td>G</td>
<td>(KER) (HD) (Q) (FMW)</td>
</tr>
</tbody>
</table>

Table 9. List of disruptive mutations for the top 25% of residues in 1a5sB at the interface with sodium ion.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>304</td>
<td>L</td>
<td>(99)M.</td>
<td>0.16</td>
<td>3/3</td>
<td>2.31</td>
</tr>
<tr>
<td>233</td>
<td>G</td>
<td>(99)R</td>
<td>0.08</td>
<td>2/2</td>
<td>4.55</td>
</tr>
<tr>
<td>256</td>
<td>E</td>
<td>(99)</td>
<td>0.11</td>
<td>1/0</td>
<td>4.99</td>
</tr>
<tr>
<td>304</td>
<td>L</td>
<td>(99)M.</td>
<td>0.16</td>
<td>1/1</td>
<td>4.21</td>
</tr>
<tr>
<td>307</td>
<td>P</td>
<td>(97)</td>
<td>0.22</td>
<td>3/3</td>
<td>4.55</td>
</tr>
</tbody>
</table>

continued in next column
Fig. 11. Residues in 1a5sB, at the interface with PLP, colored by their relative importance. The ligand (PLP) is colored green. Atoms further than 30 Å away from the geometric center of the ligand, as well as on the line of sight to the ligand were removed. (See Appendix for the coloring scheme for the protein chain 1a5sB.)

Table 11 lists the top 25% of residues at the interface with 1a5sB1. The following table (Table 12) suggests possible disruptive replacements for these residues (see Section 4.6).

Table 12. The top 25% of residues in 1a5sB at the interface with sodium ion. (Field names: res: residue number in the PDB entry; type: amino acid type; subst's: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

Figure 12 shows residues in 1a5sB colored by their importance, at the interface with 1a5sBNA2000.
Table 12.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
<td>R</td>
<td>(TD) (SYEVCAPIQ) (FMW) (N)</td>
</tr>
<tr>
<td>344</td>
<td>G</td>
<td>(KER) (FQMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>82</td>
<td>H</td>
<td>(E) (T) (Q) (D)</td>
</tr>
<tr>
<td>343</td>
<td>E</td>
<td>(FW) (H) (VCAG) (R)</td>
</tr>
<tr>
<td>80</td>
<td>L</td>
<td>(Y) (THR) (SCG) (FW)</td>
</tr>
<tr>
<td>148</td>
<td>R</td>
<td>(T) (YD) (SEVCAG) (FLMWPI)</td>
</tr>
<tr>
<td>153</td>
<td>G</td>
<td>(KR) (E) (QH) (FMWD)</td>
</tr>
<tr>
<td>54</td>
<td>G</td>
<td>(R) (KE) (FWH) (M)</td>
</tr>
<tr>
<td>345</td>
<td>I</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>381</td>
<td>D</td>
<td>(R) (H) (FW) (Y)</td>
</tr>
</tbody>
</table>

Table 12. List of disruptive mutations for the top 25% of residues in 1a5sB, that are at the interface with 1a5sB1.

Table 13.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>281</td>
<td>G</td>
<td>G(99)</td>
<td>0.11</td>
<td>1/1</td>
<td>4.97</td>
</tr>
<tr>
<td>292</td>
<td>G</td>
<td>G(99)</td>
<td>0.11</td>
<td>28/28</td>
<td>2.78</td>
</tr>
<tr>
<td>162</td>
<td>G</td>
<td>G(99)SE</td>
<td>0.15</td>
<td>5/5</td>
<td>4.64</td>
</tr>
<tr>
<td>277</td>
<td>G</td>
<td>G(99)AX</td>
<td>0.19</td>
<td>13/13</td>
<td>2.89</td>
</tr>
<tr>
<td>175</td>
<td>R</td>
<td>R(98)MA</td>
<td>0.23</td>
<td>75/7</td>
<td>3.05</td>
</tr>
<tr>
<td>14</td>
<td>G</td>
<td>G(97)</td>
<td>0.24</td>
<td>7/7</td>
<td>3.06</td>
</tr>
<tr>
<td>168</td>
<td>D</td>
<td>D(98)VE</td>
<td>0.25</td>
<td>5/2</td>
<td>4.14</td>
</tr>
</tbody>
</table>

Table 13. The top 25% of residues in 1a5sB at the interface with 1a5sA. (Field names: res: residue number in the PDB entry; type: amino acid type; subst’s: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/ bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

Table 14.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>167</td>
<td>K</td>
<td>(Y) (FW) (T) (VCAG)</td>
</tr>
<tr>
<td>281</td>
<td>G</td>
<td>(KER) (FQMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>292</td>
<td>G</td>
<td>(KER) (FQMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>162</td>
<td>G</td>
<td>(R) (K) (FWH) (E)</td>
</tr>
<tr>
<td>277</td>
<td>G</td>
<td>(R) (KE) (H) (D)</td>
</tr>
<tr>
<td>175</td>
<td>R</td>
<td>(TY) (D) (E) (S)</td>
</tr>
<tr>
<td>14</td>
<td>G</td>
<td>(KER) (FQMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>168</td>
<td>D</td>
<td>(R) (H) (FW) (KY)</td>
</tr>
</tbody>
</table>

Table 14. List of disruptive mutations for the top 25% of residues in 1a5sB, that are at the interface with 1a5sA.

Figure 14 shows residues in 1a5sB colored by their importance, at the interface with 1a5sA.

3.4.3 Possible novel functional surfaces at 25% coverage. One group of residues is conserved on the 1a5sB surface, away from (or substantially larger than) other functional sites and interfaces recognizable in PDB entry 1a5s. It is shown in Fig. 15. The right panel shows (in blue) the rest of the larger cluster this surface belongs to. The residues belonging to this surface "patch" are listed in Table 15, while Table 16 suggests possible disruptive replacements for these residues (see Section 4.6).

Table 15.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
<td>R</td>
<td>R(100)</td>
<td>0.04</td>
</tr>
<tr>
<td>84</td>
<td>G</td>
<td>G(100)</td>
<td>0.04</td>
</tr>
<tr>
<td>86</td>
<td>H</td>
<td>H(100)</td>
<td>0.04</td>
</tr>
<tr>
<td>109</td>
<td>E</td>
<td>E(100)</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 15. continued in next column

3.4.3 Possible novel functional surfaces at 25% coverage. One group of residues is conserved on the 1a5sB surface, away from (or substantially larger than) other functional sites and interfaces recognizable in PDB entry 1a5s. It is shown in Fig. 15. The right panel shows (in blue) the rest of the larger cluster this surface belongs to. The residues belonging to this surface "patch" are listed in Table 15, while Table 16 suggests possible disruptive replacements for these residues (see Section 4.6).

Figure 14 shows residues in 1a5sB colored by their importance, at the interface with 1a5sA.
**Table 15. continued**

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>299</td>
<td>S</td>
<td>S(99).</td>
<td>0.11</td>
</tr>
<tr>
<td>344</td>
<td>G</td>
<td>G(99).</td>
<td>0.11</td>
</tr>
<tr>
<td>348</td>
<td>A</td>
<td>A(99).</td>
<td>0.11</td>
</tr>
<tr>
<td>82</td>
<td>H</td>
<td>H(99)FL</td>
<td>0.12</td>
</tr>
<tr>
<td>235</td>
<td>S</td>
<td>S(99)T</td>
<td>0.12</td>
</tr>
<tr>
<td>166</td>
<td>L</td>
<td>L(99)MX</td>
<td>0.13</td>
</tr>
<tr>
<td>303</td>
<td>G</td>
<td>G(99).S</td>
<td>0.13</td>
</tr>
<tr>
<td>350</td>
<td>E</td>
<td>E(99).X</td>
<td>0.13</td>
</tr>
<tr>
<td>305</td>
<td>D</td>
<td>D(99).G</td>
<td>0.14</td>
</tr>
<tr>
<td>329</td>
<td>D</td>
<td>D(99).N</td>
<td>0.14</td>
</tr>
<tr>
<td>343</td>
<td>E</td>
<td>E(99).D</td>
<td>0.14</td>
</tr>
<tr>
<td>162</td>
<td>G</td>
<td>G(99)SEQX</td>
<td>0.15</td>
</tr>
<tr>
<td>146</td>
<td>V</td>
<td>V(98)ALXT</td>
<td>0.16</td>
</tr>
<tr>
<td>177</td>
<td>W</td>
<td>W(99)XLT</td>
<td>0.16</td>
</tr>
<tr>
<td>304</td>
<td>L</td>
<td>L(99)M.</td>
<td>0.16</td>
</tr>
<tr>
<td>80</td>
<td>L</td>
<td>L(99)MQ</td>
<td>0.17</td>
</tr>
<tr>
<td>114</td>
<td>Q</td>
<td>Q(95)M(4)E</td>
<td>0.17</td>
</tr>
<tr>
<td>148</td>
<td>R</td>
<td>R(99)KX</td>
<td>0.17</td>
</tr>
<tr>
<td>117</td>
<td>V</td>
<td>V(97)T(1)IL</td>
<td>0.18</td>
</tr>
<tr>
<td>153</td>
<td>G</td>
<td>G(99)AXS</td>
<td>0.18</td>
</tr>
<tr>
<td>159</td>
<td>V</td>
<td>V(98).A(1)XM</td>
<td>0.20</td>
</tr>
<tr>
<td>302</td>
<td>A</td>
<td>A(99).P</td>
<td>0.20</td>
</tr>
<tr>
<td>313</td>
<td>H</td>
<td>H(92)L(7).</td>
<td>0.20</td>
</tr>
<tr>
<td>382</td>
<td>K</td>
<td>K(97).(2)</td>
<td>0.21</td>
</tr>
<tr>
<td>154</td>
<td>A</td>
<td>A(97)ET(1)SXC</td>
<td>0.22</td>
</tr>
<tr>
<td>307</td>
<td>P</td>
<td>P(97)A(1)S.</td>
<td>0.22</td>
</tr>
<tr>
<td>156</td>
<td>V</td>
<td>V(2)I(97)X</td>
<td>0.23</td>
</tr>
<tr>
<td>175</td>
<td>R</td>
<td>R(98)MANKXL</td>
<td>0.23</td>
</tr>
<tr>
<td>306</td>
<td>F</td>
<td>Y(81)F(18).H</td>
<td>0.23</td>
</tr>
<tr>
<td>378</td>
<td>G</td>
<td>G(97).(2)V</td>
<td>0.23</td>
</tr>
<tr>
<td>345</td>
<td>I</td>
<td>I(99)V.</td>
<td>0.24</td>
</tr>
<tr>
<td>168</td>
<td>D</td>
<td>D(98)VESX</td>
<td>0.25</td>
</tr>
<tr>
<td>381</td>
<td>D</td>
<td>D(97).(2)XI</td>
<td>0.25</td>
</tr>
<tr>
<td>383</td>
<td>D</td>
<td>D(97).(2)X</td>
<td>0.25</td>
</tr>
</tbody>
</table>

**Table 16.**

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
<td>R</td>
<td>(TD)(SYEVCLAPIG)(FMW)(N)</td>
</tr>
<tr>
<td>84</td>
<td>G</td>
<td>(KER)(FQMWHD)(NYLPI)(SVA)</td>
</tr>
<tr>
<td>86</td>
<td>H</td>
<td>(E)(TQMD)(SNKVCLAPIG)(YR)</td>
</tr>
<tr>
<td>109</td>
<td>E</td>
<td>(FWH)(YVCARG)(T)(SNKLPI)</td>
</tr>
<tr>
<td>110</td>
<td>T</td>
<td>(KR)(FQMWHD)(NELPI)(D)</td>
</tr>
<tr>
<td>111</td>
<td>G</td>
<td>(KER)(FQMWHD)(NYLPI)(SVA)</td>
</tr>
<tr>
<td>112</td>
<td>A</td>
<td>(KYER)(QHD)(N)(FTMW)</td>
</tr>
<tr>
<td>115</td>
<td>H</td>
<td>(E)(TQMD)(SNKVCLAPIG)(YR)</td>
</tr>
<tr>
<td>138</td>
<td>D</td>
<td>(R)(FWH)(Y)(VCAG)</td>
</tr>
<tr>
<td>145</td>
<td>N</td>
<td>(Y)(FTWH)(ER)(SVCAG)</td>
</tr>
<tr>
<td>189</td>
<td>G</td>
<td>(KER)(FQMWHD)(Y)(NLPI)</td>
</tr>
<tr>
<td>232</td>
<td>G</td>
<td>(R)(K)(FWH)(EQM)</td>
</tr>
</tbody>
</table>

*continued in next column*
4 NOTES ON USING TRACE RESULTS

4.1 Coverage

Trace results are commonly expressed in terms of coverage: the residue is important if its “coverage” is small - that is if it belongs to some small top percentage of residues [100% is all of the residues in a chain], according to trace. The ET results are presented in the form of a table, usually limited to top 25% percent of residues (or to some nearby percentage), sorted by the strength of the presumed evolutionary pressure. (I.e., the smaller the coverage, the stronger the pressure on the residue.) Starting from the top of that list, mutating a couple of residues should affect the protein somehow, with the exact effects to be determined experimentally.

4.2 Known substitutions

One of the table columns is “substitutions” - other amino acid types seen at the same position in the alignment. These amino acid types may be interchangeable at that position in the protein, so if one wants to affect the protein by a point mutation, they should be avoided. For example if the substitutions are “RVK” and the original protein has an R at that position, it is advisable to try anything, but RVK. Conversely, when looking for substitutions which will not affect the protein, one may try replacing, R with K, or (perhaps more surprisingly), with V. The percentage of times the substitution appears in the alignment is given in the immediately following bracket. No percentage is given in the cases when it is smaller than 1%. This is meant to be a rough guide - due to rounding errors these percentages often do not add up to 100%.

4.3 Surface

To detect candidates for novel functional interfaces, first we look for residues that are solvent accessible (according to DSSP program) by at least 10 Å², which is roughly the area needed for one water molecule to come in the contact with the residue. Furthermore, we require that these residues form a “cluster” of residues which have neighbors within 5 Å from any of their heavy atoms.

Table 16. Disruptive mutations for the surface patch in 1a5sB.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>236</td>
<td>N</td>
<td>(Y) (FW) (TR) (VCAG)</td>
</tr>
<tr>
<td>87</td>
<td>K</td>
<td>(Y) (T) (FW) (SCG)</td>
</tr>
<tr>
<td>141</td>
<td>R</td>
<td>(T) (YD) (SECG) (VA)</td>
</tr>
<tr>
<td>142</td>
<td>Q</td>
<td>(TY) (S) (FVCAWG) (D)</td>
</tr>
<tr>
<td>233</td>
<td>C</td>
<td>(E) (D) (FMW) (YQLPHIR)</td>
</tr>
<tr>
<td>167</td>
<td>K</td>
<td>(Y) (FW) (T) (VCAG)</td>
</tr>
<tr>
<td>299</td>
<td>S</td>
<td>(KR) (FMW) (NLPI) (YE)</td>
</tr>
<tr>
<td>344</td>
<td>G</td>
<td>(KER) (FMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>348</td>
<td>A</td>
<td>(KYER) (QHD) (N) (FTMW)</td>
</tr>
<tr>
<td>82</td>
<td>H</td>
<td>(E) (T) (Q) (D)</td>
</tr>
<tr>
<td>235</td>
<td>S</td>
<td>(KR) (FMW) (NELPI) (Y)</td>
</tr>
<tr>
<td>166</td>
<td>L</td>
<td>(Y) (R) (TH) (KE)</td>
</tr>
<tr>
<td>303</td>
<td>G</td>
<td>(KR) (E) (FMW) (Q)</td>
</tr>
<tr>
<td>350</td>
<td>E</td>
<td>(FW) (Y) (VCARG) (T)</td>
</tr>
<tr>
<td>305</td>
<td>D</td>
<td>(R) (FW) (K) (VA)</td>
</tr>
<tr>
<td>329</td>
<td>D</td>
<td>(R) (FW) (YVCAG) (T)</td>
</tr>
<tr>
<td>343</td>
<td>E</td>
<td>(FW) (H) (VCAG) (R)</td>
</tr>
<tr>
<td>162</td>
<td>G</td>
<td>(R) (K) (FWH) (E)</td>
</tr>
<tr>
<td>146</td>
<td>V</td>
<td>(R) (K) (YE) (H)</td>
</tr>
<tr>
<td>177</td>
<td>W</td>
<td>(K) (E) (QR) (D)</td>
</tr>
<tr>
<td>304</td>
<td>L</td>
<td>(Y) (R) (T) (H)</td>
</tr>
<tr>
<td>80</td>
<td>L</td>
<td>(Y) (THR) (SCG) (FW)</td>
</tr>
<tr>
<td>114</td>
<td>Q</td>
<td>(Y) (H) (FW) (T)</td>
</tr>
<tr>
<td>148</td>
<td>T</td>
<td>(T) (YD) (SEVCAG) (FLWPNI)</td>
</tr>
<tr>
<td>117</td>
<td>V</td>
<td>(R) (KY) (E) (H)</td>
</tr>
<tr>
<td>153</td>
<td>G</td>
<td>(KR) (E) (QH) (FMWD)</td>
</tr>
<tr>
<td>159</td>
<td>V</td>
<td>(Y) (R) (KE) (H)</td>
</tr>
<tr>
<td>302</td>
<td>A</td>
<td>(Y) (R) (KE) (H)</td>
</tr>
<tr>
<td>313</td>
<td>H</td>
<td>(E) (T) (D) (Q)</td>
</tr>
<tr>
<td>382</td>
<td>K</td>
<td>(Y) (FTW) (SVCA) (HD)</td>
</tr>
<tr>
<td>154</td>
<td>A</td>
<td>(R) (K) (Y) (EH)</td>
</tr>
<tr>
<td>307</td>
<td>P</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>156</td>
<td>V</td>
<td>(YR) (KE) (H) (QD)</td>
</tr>
<tr>
<td>175</td>
<td>R</td>
<td>(TY) (D) (E) (S)</td>
</tr>
<tr>
<td>306</td>
<td>F</td>
<td>(KE) (Q) (D) (T)</td>
</tr>
<tr>
<td>378</td>
<td>G</td>
<td>(KER) (HD) (Q) (FMW)</td>
</tr>
<tr>
<td>345</td>
<td>I</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>168</td>
<td>D</td>
<td>(R) (H) (FW) (KY)</td>
</tr>
<tr>
<td>381</td>
<td>D</td>
<td>(R) (H) (FW) (Y)</td>
</tr>
<tr>
<td>383</td>
<td>D</td>
<td>(R) (FWH) (Y) (VCAG)</td>
</tr>
</tbody>
</table>

Table 16. continued

4.4 Number of contacts

Another column worth noting is denoted “noc/bb”: it tells the number of contacts heavy atoms of the residue in question make across the interface, as well as how many of them are realized through the backbone atoms (if all or most contacts are through the backbone, mutation presumably won’t have strong impact). Two heavy atoms are considered to be “in contact” if their centers are closer than 5 Å.

4.5 Annotation

If the residue annotation is available (either from the pdb file or from other sources), another column, with the header “annotation” appears. Annotations carried over from PDB are the following: site (indicating existence of related site record in PDB ), S-S (disulfide bond forming residue), hb (hydrogen bond forming residue, jb (james bond forming residue), and sb (for salt bridge forming residue).

4.6 Mutation suggestions

Mutation suggestions are completely heuristic and based on complementarity with the substitutions found in the alignment. Note that they are meant to be disruptive to the interaction of the protein with its ligand. The attempt is made to complement the following properties: small [AVGSTC], medium [LPNQDEM1K], large [WFYHR], hydrophobic [LPVAMWF1], polar [GTYC]; positively [KHR], or negatively [DE] charged, aromatic [WFYH], long alkyl chain [EKIQM], OH-group possession [SDETY], and NH2 group possession [NQRK]. The suggestions are listed according to how different they appear to be from the original amino acid.
The colors used to distinguish the residues by the estimated evolutionary pressure they experience can be seen in Fig. 16.

5.3 Credits

5.3.1 Alistat Alistat reads a multiple sequence alignment from the file and shows a number of simple statistics about it. These statistics include the format, the number of sequences, the total number of residues, the average and range of the sequence lengths, and the alignment length (e.g. including gap characters). Also shown are some percent identities. A percent pairwise identity is defined as (idents / MIN(len1, len2)) where idents is the number of exact identities and len1, len2 are the unaligned lengths of the two sequences. The "average percent identity", "most related pair", and "most unrelated pair" of the alignment are the average, maximum, and minimum of all (N(N-1)/2 pairs, respectively. The "most distant seq" is calculated by finding the maximum pairwise identity (best relative) for all N sequences, then finding the minimum of these N numbers (hence, the most outlying sequence). alistat is copyrighted by HHMI/Washington University School of Medicine, 1992-2001, and freely distributed under the GNU General Public License.

5.3.2 CE To map ligand binding sites from different source structures, report_maker uses the CE program: http://cl.sdsc.edu/. Shindyalov IN, Bourne PE (1998) "Protein structure alignment by incremental combinatorial extension (CE) of the optimal path". Protein Engineering 11(9) 739-747.

5.3.3 DSSP In this work a residue is considered solvent accessible if the DSSP program finds it exposed to water by at least 10 Å, which is roughly the area needed for one water molecule to come in contact with the residue. DSSP is copyrighted by W. Kabsch, C. Sander and MPI-MF, 1983, 1985, 1988, 1994, 1995, CMBI version by Elmar.Krieger@cmbi.kun.nl November 18, 2002, http://www.cmbi.kun.nl/gv/dssp/descrip.html.


5.3.5 LaTeX The text for this report was processed using LaTeX: Leslie Lamport, "LaTeX: A Document Preparation System Addison-Wesley," Reading, Mass. (1986).


5.3.7 Pymol The figures in this report were produced using Pymol. The scripts can be found in the attachment. Pymol is an open-source application copyrighted by DeLano Scientific LLC (2005). For more information about Pymol see http://pymol.sourceforge.net/. (Note for Windows

Fig. 16. Coloring scheme used to color residues by their relative importance.

The text for this report was processed using LaTeX; Leslie Lamport, “LaTeX: A Document Preparation System Addison-Wesley,” Reading, Mass. (1986).
users: the attached package needs to be unzipped for Pymol to read the scripts and launch the viewer.)

5.4 Note about ET Viewer
Dan Morgan from the Lichtarge lab has developed a visualization tool specifically for viewing trace results. If you are interested, please visit:

http://mammoth.bcm.tmc.edu/traceview/

The viewer is self-unpacking and self-installing. Input files to be used with ETV (extension .etvx) can be found in the attachment to the main report.

5.5 Citing this work


5.6 About report_maker
report_maker was written in 2006 by Ivana Mihalek. The 1D ranking visualization program was written by Ivica Reš. report_maker is copyrighted by Lichtarge Lab, Baylor College of Medicine, Houston.

5.7 Attachments
The following files should accompany this report:

- 1a5sA.complex.pdb - coordinates of 1a5sA with all of its interacting partners
- 1a5sA.etvx - ET viewer input file for 1a5sA
- 1a5sA.cluster_report.summary - Cluster report summary for 1a5sA
- 1a5sA.ranks - Ranks file in sequence order for 1a5sA
- 1a5sA.clusters - Cluster descriptions for 1a5sA
- 1a5sA.msf - the multiple sequence alignment used for the chain 1a5sA
- 1a5sA.descr - description of sequences used in 1a5sA msf
- 1a5sA.ranks_sorted - full listing of residues and their ranking for 1a5sA
- 1a5sA.1a5sAIFIP270.if.pml - Pymol script for Figure 5
- 1a5sA.cbcvg - used by other 1a5sA – related pymol scripts
- 1a5sA.1a5sB.if.pml - Pymol script for Figure 6
- 1a5sB.complex.pdb - coordinates of 1a5sB with all of its interacting partners
- 1a5sB.etvx - ET viewer input file for 1a5sB
- 1a5sB.cluster_report.summary - Cluster report summary for 1a5sB
- 1a5sB.ranks - Ranks file in sequence order for 1a5sB
- 1a5sB.clusters - Cluster descriptions for 1a5sB
- 1a5sB.msf - the multiple sequence alignment used for the chain 1a5sB
- 1a5sB.descr - description of sequences used in 1a5sB msf
- 1a5sB.ranks_sorted - full listing of residues and their ranking for 1a5sB
- 1a5sB.1a5sCPLP901.if.pml - Pymol script for Figure 11
- 1a5sB.cbcvg - used by other 1a5sB – related pymol scripts
- 1a5sB.1a5sBNA2000.if.pml - Pymol script for Figure 12
- 1a5sB.1a5sB1.if.pml - Pymol script for Figure 13
- 1a5sB.1a5sA.if.pml - Pymol script for Figure 14